Educational Leadership / RESEARCH SUPPLEMENT

EDITED BY ASCD RESEARCH COUNCIL • JAMES RATHS, CHAIRMAN

Volume 2 **May 1969** Number 5

CONTENTS

The Effects of Set Induction Upon Pupil Achievement, Retention, and Assessment of Effective Teaching in a Unit on Respiration in the BSCS Curricula

***ROBERT F. SCHUCK* ***

SEVERAL psychologists have contended that one of the essential conditions in learning is that the learner be actively engaged in the process in order to maximize the possibility of new behavioral responses. Studies and related readings on *set* indicate that this is a powerful variable in determining the kinds of learning that will occur in the classroom and that the teacher is an instrumental agent in arousing or inducing the pupil's set toward learning. An experiment on the application of set in the student teaching environment conducted by Wittrock at the University of California at Los Angeles revealed that set appeared to have relevance to pupil learning. He reported that pupils taught by student teachers who were told that their grade in student teaching depended upon the achievement gains made by their students, made significant gains in achievement over pupils taught by student teachers in a control group who were not given such a set (.05 level). (6:179)

In a study at Stanford University, Aubertine outlined procedures used to train teachers in the skill of incorporating set induction techniques in their instructional strategies. He reported that teachers who did incorporate these procedures into their instructional strategies were perceived as more effective

* *Robert F. Schuck, Assistant Professor of Education, School of Education, University of Connecticut, Storrs*

by their pupils than were a control group of teachers who did not receive training in set induction procedures (.05 level). (1:69)

Statement of the Problem

This study combines one of the latest developments in the substantive and environmental dimensions of biology education (the BSCS) with one of the latest developments in the behavioral dimensions of the educative process (set induction procedures) to assess the total impact of this combination upon the learner. Specifically, this study seeks to answer the following questions:

1. Will ninth grade pupils taught by teachers trained in set induction techniques make significant gains in achievement over those taught by teachers who have not been so trained?

2. Is there a significant difference in the ratings ninth grade pupils give teachers trained in set induction procedures regarding teacher effectiveness over those given to teachers not trained in set induction procedures?

3. Will ninth grade pupils taught by teachers trained in set induction techniques retain the material taught to a significantly greater degree than pupils who have been taught by teachers who have not been so trained?

4. Does a significant correlation exist between the ratings ninth grade pupils give their teachers on the Teacher Demonstration Rating Form and the achievement and retention these pupils register on a pretest—post-test design?

If significant differences can be shown in the areas of pupil assessment of effective teaching, pupil achievement, and pupil retention between groups using the BSCS materials alone and those using the BSCS materials in combination with teacher use of set induction techniques, then an important advancement in our knowledge of the educative process has been made. Such data would provide a strong argument for greater effort in combining the curriculum developments and their advances in the substantive and environmental dimensions with elements of the behavioral dimension as educators seek to enhance the quality of their instructional programs.

If, as a result of this study, no significant differences can be found, then the present strategy of curriculum development would gain some support.

In summary, this investigation was designed to answer the question: Is it important that educators take into consideration the behavioral dimension of the educative process in planning programs for biology education at the secondary level or are the present programs with their emphasis on the environmental and substantive dimension adequate for the attainment of desired learning goals?

For the purpose of this study, the following definitions were employed:

Set: A set is a cognitive process activated by a stimulus or stimuli perceived by a person in environmental situations, determining how one is predisposed to respond to what is attended to in a given situation.

Set Induction: Inducing a learning set is the initial instruction act on the part of the teacher for the purpose of establishing a frame of reference deliberately designed to facilitate the creation of a communicative link between the experimental field of the pupils and the desired behavioral goals of the learning experience (the lesson).

The purpose of the set induction procedure is to focus pupil attention on some commonly known experimental referent (*orientation*) which becomes the vehicle by which the teacher makes the passage from the known to new material (*transition*) and builds continuity from lesson to lesson. The induced set lends meaning to new material through the use of analogy rather than by simple association (*operation*) and encourages pupil involvement in the lesson as judged by the teacher (*evaluation*).

Example

In a biology class on an introductory unit on "The Process of Photosynthesis," the teacher might induce a set in the first lesson in the following manner:

1. *Orientation.* The teacher begins the lesson by exhibiting a bottle of milk for the

purpose of generating a discussion of the process of milk production by cows.

2. *Transition.* The teacher uses the cow as an analogue to a manufacturing plant in which it consumes raw material (hay and water), adds its own enzymes and energy to produce a product and a by-product (milk and manure). This process may be simply illustrated on the blackboard in the following way: raw material (hay and water) to cow (enzymes and energy) to product and by-product (milk and manure).

3. *Operation.* The teacher introduces the process of photosynthesis by relating it to the above analogue in the following manner: raw material (water and carbon dioxide) to process (sunlight and chlorophyll) to product and by-product (sugar and oxygen). This could be illustrated with overlays on an overhead projector.

4. *Evaluation.* The teacher then seeks pupil comprehension through asking general questions about the process of photosynthesis. When the teacher believes that the pupils have grasped the concept, he proceeds into the main body of the lesson with more difficult material and eventually to the Krebs cycle.

In the course of the experiment several additional set induction strategies were developed. Each utilized the format explained in the preceding paragraphs; however, they dealt with the substantive material covered in the respective BSCS unit on respiration.

Experimental Design

The experimental design employed in this study was the Pretest—Post-test Control Group design. This design was incorporated in the following manner:

Pupil Sample

A volunteer population of ninth grade pupils was solicited from E. O. Smith High School located in Storrs, Connecticut. The total population of pupils used was 180. These pupils were then randomly divided into 18 groups numbering 10 pupils per group. This resulted in the establishment of a total of 18 separate groups numbering 10 pupils per group.

Teacher Sample

The teacher population consisted of 18 education majors who had not taken any methods courses, nor had they been exposed to the student teaching experience. These teachers were volunteers who responded to a letter requesting their participation in the study. This letter was sent to each of the students in the College of Education.

Each of the 18 teachers participating in the study was assigned at random to one of the pupil groups already established. In turn, each of the groups, with the teacher assigned to that group, was randomly assigned to either the experimental or control procedure.

Curriculum Assignment

The BSCS Curriculum to be used by each teacher in instructing the pupils assigned to him was determined at random. This resulted in having three teachers instruct 10 pupils each in the BSCS Blue, Green, and Yellow Versions in both the experimental and control groups.

A unit on respiration was employed because of its noncontroversial nature and the fact that the 12 instructional periods employed in the design exceeded the instructional time allotment recommended by the BSCS program (4:61).

Summary of Design

The actual implementation of the design chosen resulted in a high degree of random assignment. All of the following assignments were accomplished by employing random procedures.

1. Pupils to groups of 10 each

2. Teachers to each of the pupil groups formed

3. Pupil-teacher group to be used in the experimental or the control procedure

4. The unit on respiration in the BSCS Curriculum to be used to each teacher-pupil group.

Experimental Treatment

All participating teachers attended a two-hour meeting one week prior to the first

week of instruction. At this meeting their assignment to pupil groups and the BSCS Curricula to be used were made known. Teachers were then assigned to teams within the experimental and control groups by the BSCS Curricula used. To eliminate the threat of contamination, no team contained teachers from both the experimental and control groups. The formation of these teams, however, made the opportunity available for teachers using the same BSCS Curriculum to pool information concerning their instructional strategies.

This first meeting was the only time the entire control group of teachers met together. The experimental group of teachers held three additional meetings on the Saturday prior to each week of instruction.

The purpose of these meetings was to give instruction to the experimental groups of teachers on set induction techniques and how these could be incorporated into their instructional strategies. These meetings lasted two hours each and gave a total training of six additional hours to the experimental group. During these training sessions the experimental group of teachers was exposed to the set induction training procedures.

Once these procedures were thoroughly covered, the remaining time was spent in reviewing and discussing a few samples of sets employed by teachers in studies conducted at Stanford to illustrate more effectively for the group an operational instructional strategy employing set induction techniques. Members of the group were then dismissed with the request that they attempt to incorporate set induction techniques in the formation of their instructional strategies and that they communicate only with the members of their three-man team.

The second and third sessions were held on the morning of the Saturday falling between the first and second and the second and third week instructional periods. During these sessions questions regarding the use of sets during the previous week were answered. A review of the set induction procedures was then undertaken. The final activity was an analysis of several sets submitted by members of the group during the previous week. These sets were dittoed and distributed to the experimental group of teachers. The total time spent in training the experimental group of teachers in set induction techniques was six hours.

The control group of teachers received no training beyond the initial meeting, in which all 18 teachers were given the material for their respective BSCS Curriculum. The control group of teachers then depended upon only this material to form their instructional strategies.

Methods

The data necessary in this study were collected through employment of both pretests and post-tests. Two tests were used: the Teacher Demonstration Rating Form developed at Stanford University and the Achievement Test constructed at Arizona State University especially for this study.

The instructional period for this study consisted of M,W,F, for four consecutive weeks. During these weeks the classes were conducted after school from 2:20 to 3:30 p.m. Monday through Friday. The total time involvement was 12 hours. This exceeds the recommended time given the unit on respiration by the BSCS program (4:61).

Table 1 illustrates the instructional and testing pattern implemented in this study. The X's represent the testing period for the Achievement Test, the O's are pure instructional periods, and the Z's are testing periods for the T.D.R. The subscript 1 indicates a pretest situation, and subscript 2 indicates a post-test situation.

Week	Mon.	Wed.	Fri.
1	X_1	Z_1	O
2	O	O	O
3	O	O	O
4	O	Z_2	X_2
8	Retention Test		

Table 1. The Schedule of Instruction and Data Collection

The Achievement Test was administered at the beginning of the period. The T.D.R.

was always administered in the final 15 minutes of the period because its purpose was to evaluate the effectiveness of the teacher instruction. The Retention Test was administered four weeks following the completion of the instruction.

The Instrument Used To Collect Data

The data analyzed to test the hypotheses developed in this study were collected through employment of two separate measurement instruments—the Teacher Demonstration Rating Form to measure pupil perception of effective teaching and the Achievement Test designed to measure pupil achievement and retention in a unit on respiration in the BSCS Curricula.

Validity of the Achievement Test

Content validity was established by requesting the 12 BSCS teachers, four teaching in each of three curricula, to submit 20 multiple choice items which they felt would cover the unit on respiration adequately. These questions were then coded for the curriculum the teacher taught and the resulting test of 240 items was constructed. The test was then determined to have content validity due to the fact that a trained group of teachers had determined the questions as adequate for the purpose of measuring achievement in the unit on respiration in their respective BSCS program.

The next step was to establish item validity. The procedure used was that suggested by Henry Garrett (3:351). The method suggested for determining the maximum item validity was to develop a test consisting of items with:

1. Difficulty indices of more than .20
2. Discrimination indices between .50 and .80.

The original test composed of 240 items was then administered to a sample of ninth grade students in the greater Phoenix area. These students had already been exposed to the unit on respiration so that it was possible to select 40 students from each of three BSCS Curricula. An item analysis was made of the test responses and those items not satisfying Garrett's criteria were eliminated. The resulting number of items remaining was 113.

These items were then combined, forming a new test. This test was administered to a new student population consisting of ninth grade pupils having taken the unit on respiration in the BSCS program, N=126. These were students from high schools independent of the first group employed and also independent of the population utilized in the study.

The items not fitting the categories were eliminated and the remaining items were distributed as follows: 15 from the Blue Version, 27 from the Yellow Version, and 19 from the Green Version, making a total of 61 items.

All items were then placed into separate pools by BSCS Version, and a random selection of 13 items was drawn from each, yielding a test of 39 items. The remaining 22 items were then mixed, forming a population from which the remaining one item was drawn, so that a 40-item test was constructed. The item selected at random was from the Yellow Version, and the construction of the resulting 40-item test was Yellow—14 items, Blue—13 items, and Green—13 items.

Since the available population of ninth grade students having taken BSCS biology was nearly exhausted, a smaller administration of the 40-item test was made. The N here was 30, 10 from each version, and an analysis of the data collected showed all of the test items fell within the criteria listed by Garrett.

Thus, three pilot studies of the achievement test were made. The resulting test was determined to be a valid one since it was composed of items all found to lie within the limits of discrimination and difficulty established by Garrett for valid test items.

After the test was administered to the experimental and control populations as a post-test, another check was made of the items. Only one of the items fell outside of the criteria, that item having a discrimination index of .83. Since this deviation was small, the test was once again determined to be a valid one.

RELIABILITY

Reliability was established through use of the Kuder Richardson 20. The reliability figures on the pilot studies reported were .58, .72, and .78 respectively. The reliability coefficient found by use of the K-R 20 on the experimental post-test given to all students in the study was .86.

Results

The experimental group made a total mean gain from pre- to post-test scores on both the Achievement Test and the Teacher Demonstration Rating Form. These overall mean gains were found to be statistically significant at less than .01 level of confidence using the Fisher F test for the Analysis of Variance. In relation to the hypotheses tested, the following was found:

Hypothesis 1

There is no significant difference in the ratings pupils give student teachers on the Teacher Demonstration Rating Form in the experimental vs. the control group.

FINDINGS

Using the raw scores teachers received on the T.D.R., an analysis of variance was applied to determine the F value of the difference. The data cited in Table 2 show the test of significance of these T.D.R. scores between experimental and control groups. The F value for the difference was determined to be 84.37 and the F table indicated that 6.78 (an approximate interpolated value) was needed to be significant at the .01 level of confidence. The null hypothesis of no significant difference between the T.D.R. scores of teachers in the experimental vs. the control group was rejected.

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Sum of Squares
Between	1.	396.03	1308.65
Within	178.	912.62	15.51
	179.	1308.65	

$$F1, 178 = \frac{1308.65}{15.51} = 84.37$$

(significant at the .01 level of confidence)

Table 2. A Comparison of the T.D.R. Scores Made by Students in the Experimental and Control Groups

Hypothesis 2

There is no significant difference in achievement on a biological unit in respiration in the BSCS program as measured through a pretest—post-test design between the experimental and control group of high school pupils.

FINDINGS

Using the raw scores pupils received on the Achievement Test, an analysis of variance was applied to determine the F value of the difference. The data cited in Table 3 show the test of significance with the Achievement Test between experimental and control groups. The F value for the difference was determined to be 33.47 and the F table indicated that 6.78 (an approximate interpolated value) was needed to be significant at the .01 level of confidence. The null hypothesis of no significant difference in the achievement between the pupils in the experimental and control group was therefore rejected.

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Sum of Squares
Between	1.	681.63	681.63
Within	178.	1202.03	20.36
	179.	1883.66	

$$F1, 178 = \frac{681.63}{20.36} = 33.47$$

(significant at the .01 level of confidence)

Table 3. A Comparison of the Achievement Test Scores Made by Students in the Experimental and Control Groups

Hypothesis 3

There is no significant difference in the retention of information on a biological unit in respiration in the BSCS program as measured through a pretest—post-test design between the experimental and control groups of high school pupils.

FINDINGS

Using the raw scores pupils received on the Achievement Test, an analysis of variance was applied to determine the F value of the difference. The data cited in Table 4 show the test of significance with the Achievement Test between the experimental and control groups. The F value for the difference was determined to be 35.30 and the F table indicated that 6.78 (an approximate interpolated value) was needed to be significant at the .01 level of confidence. The null hypothesis of no significant difference in the achievement between the pupils in the experimental and the control group was therefore rejected.

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Sum of Squares
Between	1.	880.20	880.20
Within	178.	1471.38	24.93
	179.	2351.58	

$$F_{1,178} = \frac{880.20}{24.93} = 35.30$$

Table 4. A Comparison of the Retention Test Scores Made by Students in the Experimental and Control Groups

Hypothesis 4

There is no significant correlation between the ratings high school pupils give their teachers on the Teacher Demonstration Rating Form and the achievement these pupils register on the Achievement Test in a pretest—post-test design.

FINDINGS

The data cited in Table 5 show the correlation value found. The correlation value necessary for significance in the control and experimental group when the N=90 for each was .27 for significance for the .01 level of confidence. Since the value found for both groups was greater than .27, the correlations between those students rating their teachers as most effective and their achievement was determined to be significant at the .01 level of confidence.

An analysis of the data obtained was made to determine if a significant correlation existed between the score a student gave his teacher on the T.D.R. and the amount of gain made in achievement by the student. It must be remembered that on the T.D.R., the lower the score received, the more effective the teacher is perceived to be.

Experimental Group	Control Group
−.68	−.58
(significant at .01 level of confidence)	

Table 5. Overall Correlation of Student Perception of Effective Teaching with Student Achievement

Hypothesis 5

There is no significant correlation between the ratings high school pupils give their teachers on the Teacher Demonstration Rating Form and the scores of these pupils registered on an Achievement Test.

Experimental Group	Control Group
−.54	−.32
(significant at .01 level of confidence)	

Table 6. Overall Correlation of Student Perception of Effective Teaching with Student Retention

FINDINGS

The data cited in Table 6 show the correlation value found. The correlation value necessary for significance in the control and experimental groups when the N=90 for each was .27 for significance of the .01 level of confidence. Since the value found for both groups was greater than .27, the correlations between those students rating their teachers are most effective and their degree of retention was determined to be significant at the .01 level of confidence.

Discussion

The data collected in this investigation indicated that pupils of teachers trained in the deliberate skill of set induction did achieve significantly higher on the Achievement Test and Retention Test employed, and they did perceive their teachers to be significantly more effective as measured by the Teacher

Demonstration Rating Form beyond the .01 level of confidence. Furthermore, a correlation significant at the .01 level of confidence was found in both groups between the amount of gain in achievement made by the student, the degree of retention, and the student's perceived effectiveness of his teacher.

A closer examination of the data revealed some interesting information. When the mean scores of the Pretest Achievement Test of the experimental and control groups were compared, a difference of .6 of one point in favor of the experimental group was found. When a comparison was made of the mean scores of the Pretest Teacher Demonstration Rating Form between the experimental and control, a difference of .9 of one point was found in favor of the experimental group.

The high degree of randomization employed in the design of this study appeared to be a prime factor in the production of the close proximity of these mean scores. The relatively low mean scores of both groups on the Teacher Demonstration Rating Form Pretest was seen as an indicator that the Hawthorne Effect was in operation within both groups. Previous use of this instrument had resulted in mean scores from one to three points higher. It must be remembered that on this instrument the more effective a teacher was perceived as being, the lower the score he would be given. Hence, the conclusion was reached that the low means obtained were the result of the Hawthorne Effect in operation.

The Post-test data in the area of achievement revealed a difference between the experimental and the control group mean scores of 5.1 points in favor of the experimental group. This indicated that the experimental group began with .6 of a point advantage, and upon the Post-test obtained a 5.1 difference making a total net gain of 4.5 points, a difference as was noted earlier, which was significant beyond the .01 level of confidence.

An analysis of the Teacher Demonstration Rating Form Pretest scores led to a belief that a statistical regression toward the mean would operate. If significant differences were to result on the Post-test, it was felt that these differences would favor the group which regressed the least. In other words, the group which held its original low scores would be judged most effective. The data collected from the Post-test of the Teacher Demonstration Rating Form indicated that the mean score of the control group had moved from 13.4 to 16.3. This shift indicated a regression of 2.9 points. On the other hand, the mean score of the experimental group shifted from 13.3 to 11.7. This shift totaled 1.6 points. This indicated a gain of 1.6 points since on this instrument a low rating indicated a higher perception of teacher effectiveness. The data collected revealed that rather than demonstrating a regression as was thought, the experimental group actually gained a total of 1.6 points. This resulted in a final difference between the mean scores of the experimental and the control groups of 4.5 points in favor of the experimental group. The difference was found to be significant beyond the .01 level of confidence.

The findings of this study may be summarized as follows:

1. Teachers within the experimental group received statistically significant higher ratings on the Teacher Demonstration Rating Form (.01 level of confidence) than the teachers in the control group.

2. Pupils within the experimental group made statistically significant mean gains on the Achievement Test (.01 level) over those in the control group.

3. Pupils within the experimental group made significant mean gains in retention (.01 level) over those in the control group.

4. The correlations between those students recording high achievement gains and their ratings of teacher effectiveness were statistically significant (.01 level of confidence) within both the experimental and control groups.

Conclusions

1. Teachers who are trained in the deliberate use of set induction techniques in their instructional strategies will be viewed by their pupils as being significantly more effective.

2. Pupils taught by teachers trained in the deliberate use of set induction techniques will achieve significantly higher than those ex-

posed to teachers not trained in this instructional skill.

3. Pupils who view their teachers as effective tend to make greater gains in achievement and to retain better the material taught.

References

1. Horace E. Aubertine. "An Experiment in Set Induction Processes and Its Application in Teaching." Ed.D. dissertation. Stanford: Stanford University, 1964.

2. Lawrence Downy. *The Secondary Phase of Education*. New York: Blaisdell Publishing Company, 1965.

3. Henry E. Garrett. *Statistics in Education and Psychology*. New York: Longmans, Green and Co., 1950.

4. John A. Moore. *Teacher's Manual: Biological Sciences: An Inquiry Into Life*. New York: Harcourt, Brace & World, Inc., 1964.

5. Robert F. Schuck. "An Investigation To Determine the Effects of Set Induction Upon the Achievement of Ninth Grade Pupils and Their Perceptions of Teacher Effectiveness in a Unit on Respiration in the BSCS Curricula." Ed.D. dissertation. Tempe: Arizona State University, 1967.

6. M. C. Wittrock. "Set Applied to Student Teaching." *Journal of Educational Psychology* 53: 175-80; August 1962. □

Notice to Contributors

WITH this issue, the editorship of the RESEARCH SUPPLEMENT changes hands. The new chairman of the ASCD Research Council, Professor Frederick A. Rodgers of New York University, will serve as editor of the RESEARCH SUPPLEMENT for the next three years. Authors are invited to submit their manuscripts to Professor Frederick A. Rodgers, Chairman, ASCD Research Council, Association for Supervision and Curriculum Development, 1201 Sixteenth Street, N.W., Washington, D.C. 20036. c/o Dr. Robert R. Leeper.

This RESEARCH SUPPLEMENT is *not* designed for publishing reviews of research, discussions of research issues, calls for needed research analysis, or analysis of widely quoted research studies. *It has been established for the reporting of data.* The criteria for judging which articles are to be included in the limited space are listed in a hierarchical sequence as follows:

1. The manuscript must report data. Included in the article must be some evidence to support the reliability of the measures used in the study.

2. The article should concern itself with the behavior of teachers (or their surrogates) and that of students as dependent variables. Behavior is taken to mean achievement scores, responses to questionnaires, etc.

3. The article should present a discussion of the results in such a manner that the meaning of the research is clear to readers. Some suggestions to meet this criterion include: a discussion of threats to the validity of the study's conclusion; an unambiguous definition of the independent variable; a distinction between the findings (data) of a study and the conclusion pertaining to the research hypotheses; a distinction between testing research hypotheses grounded in theoretical frameworks and answering research questions for which there exists no known theoretical base; and finally establishing a basis for qualified conclusions.

Manuscripts are welcome in all lengths, from 500 words to 8800 words, typed, double-spaced. (Most articles in *Educational Leadership* are roughly 1500 words in length.) The authors of manuscripts are invited to use the style exemplified by the regular articles found in *Educational Leadership* of citing references at the end of every chapter in alphabetical order. All manuscripts will be submitted to panels of referees drawn from the membership of the ASCD Research Council and prompt decisions will be made regarding the publication of manuscripts. It is hoped that researchers in the field will consider submitting reports of their work to the RESEARCH SUPPLEMENT and that leaders in the field of curriculum development and supervision will direct their attention to the researches reported herein. □